NETWORK SECURITY DEVICE AND METHOD

RELATED APPLICATION

[0001] This application claims priority to Provisional Application Serial No. 60xxxxxx, filed October 19, 2001.

TECHNICAL FIELD

[0002] This invention relates to a technique for achieving a high level of physical security in a network security device such as would be used with a portable computer, a computer terminal or a Personal Digital Assistant (PDA) to connect to a network and obtain secure service from that network.

BACKGROUND ART

[0003] The power of a computing device, such as a personal computer, data terminal or even a Personal Data Assistant (collectively referred hereinafter as "a network peripheral device") improves dramatically when such device is connected to other devices across a network to allow information sharing. Such a network may take the form of a simple Local Area Network (LAN), Wide Area Network, Corporate Intranet, the Intranet or combination of such networks. In many instances, the services, resources and/or data accessed or transmitted through this network are sensitive in that a breech of authenticity or privacy of the services, resources or information would have economic or other undesirable consequences for the users of the network. .

[0004] Security is achieved by the use of a combination of software and hardware measures. Software employing a variety of cryptographic techniques is used to encrypt and/or authenticate the information exchanged through the network while hardware-based physical security measures guarantee that the cryptographic keys and the software using these keys remain uncorrupted, private and trustworthy. The software and cryptographic techniques used depend on the services, resources and information accessed through the network; for example, a network security device that supports

Virtual Private Networking (VPN) functionality will have software that implements IPsec, PPTP or some other VPN protocol. This software will use cryptographic keys in the way specified by the VPN protocol in use to encrypt and/or authenticate all information flowing to and fro the network.

[0005] Physical security can be achieved in different ways. Two approaches to physical security are common: physical access control and tamper-proofing. In the first approach, no specific physical security measures are included in the device; the physical security depends entirely on the fact that only authorized and trustworthy users have physical access to the device. In the second approach, the casing of the device is hardened to make its penetration difficult and detectors are placed inside the device to detect any attempt to break through the casing; if a penetration attempt is detected, the device erases all sensitive information from its memory and renders itself useless. The level of security afforded by the first approach depends on the inaccessibility of the device and is limited by the fact that there will be no way to detect a compromise of the device if the physical access controls fail. In most settings the second approach affords a much higher level of security. However, tamper-proofing by itself is not enough to guard against substitution attacks. In a substitution attack, the attacker replaces the security device of the user by another similar device that was prepared specially so that it uses keys known to the attacker, thereby nullifying the security provided by the device for the user. Tamper-proofing a device is also expensive: the device has to be augmented to include intrusion detectors, circuitry that continuously monitors the detectors and some power source to keep the intrusion detection system active when the device is not in use.

[0006] Thus, a need exists for a physical security mechanism that guarantees the integrity of software and keys used by the device and that protects against substitution attacks while keeping the cost of the security measures low.

BRIEF SUMMARY OF THE INVENTION

[0006] Briefly, in accordance with a preferred embodiment, the present invention provides a combination of physical security mechanisms and restrictions on the software placed in the device that together provide a high level of security, protecting the device's

user (or users) against tampering of the device and against substitution attacks. Because of the restrictions on the software, not all network security devices can benefit from this invention; only those whose software can be modified to fit the imposed restrictions. In practice this does not restrict the types of services that can be offered by the device, only the specific cryptographic protocols that can be used to secure these services. For example, a VPN card implementing PPTP cannot make use of this invention because PPTP does not have the 'perfect forward secrecy' property. On the other hand, a VPN card implementing IPsec can. The security mechanism of the invention includes at least one immutable memory element (e.g., a read-only memory element) that contains information that remains immutable (unchanged) prior to and after each session (except for any upgrades). In practice, the immutable memory element contains security application code that "boot straps" (initiates the operation of) the security mechanism itself as well as initiating execution of application code that provides the security services (i.e., user and security mechanism authentication). The security mechanism also includes a persistent memory element that contains files that may undergo a change between sessions. For example, the persistent memory element may contain configuration information that permits the user to gain network access in different environments. Lastly, the security mechanism includes a volatile memory element for retaining data for only the length of a current session. For example, the volatile memory element typically contains critical security data (e.g., a user password or session specific cryptographic keys) to permit connection to the network as well as provide authentication data that authenticates the user and the security mechanism itself. At the end of the session, all of the information in the volatile memory element is erased, thereby preventing re-use of such information by unauthorized users. A tamper-evident enclosure contains the memory elements. The tamper-evident enclosure, when tampered with, will reflect such tampering, thereby allowing the user to know if an attempt was made to physically compromise the security mechanism.

[0007] The security mechanism of the invention affords a high level of security if the software of the device can be made to meet the requirements for 'perfect forward security' and if the device obtains all security critical data from its user at the beginning of each session. In the context of this invention, we define perfect forward security as the

property of software whereby a future compromise of the device will not compromise past or present sessions protected by that device. At the beginning of a session, the security mechanism executes a 'key exchange' with the remote gateway. In this exchange, a session key is generated at random, encrypted using the device's private key and sent to the gateway. An attacker who intercepts this encrypted message and later gets access to the device could extract the device's private key and use that to decrypt the session key. In this manner, the attacker breaks the security of a past session. Perfect forward secrecy refers to esoteric cryptographic techniques that render this type of attack impossible. This definition is an extension of the concept of 'perfect forward secrecy' that is a property of cryptographic key exchange protocols that has been much discussed in the cryptographic research community. A consequence of the perfect forward security requirement for the device's software is that any key exchange protocol it uses must have the perfect forward secrecy property. As discussed, the volatile memory that holds the authentication information for the current session is erased at the end of a current session, preventing its re-use. Thus, if someone were to misappropriate the security mechanism, no authentication information remains to allow for unauthorized network entry and no information remains that could be used to decrypt a past session. Moreover, since a tamper-evident enclosure surrounds the various memory elements of the security mechanism, any attempt to physically gain access would become apparent to the legitimate holder of the security mechanism. The security critical data that the device obtains from the user at the beginning of a session must be sufficient to unambiguously determine the security services expected by the user.

[0008] The perfect forward security requirement guarantees that a compromise of the device will not compromise the security of past sessions. The tamper-evident properties of the enclosure guarantee that the user will not entrust sensitive information to a device that was compromised. Finally, the requirement that the device collects security critical data at the beginning of each session guarantees that an uncompromised device will provide the expected security services thereby guarding against substitution attacks.

BRIEF DESCRIPTION OF THE DRAWING

[0007] FIGURE 1 illustrates partially cut-away perspective view of a security device in accordance with a preferred embodiment of the present invention;

DETAILED DESCRIPTION

[0010] FIGURE 1 illustrates a security mechanism 10 in accordance with a preferred embodiment of the invention for permitting a user (not shown) to establish a secure communications session between a network peripheral device 12 and a communications network 14 via a security gateway 15. The network peripheral device 12 can take the form of a computer terminal, personal computer or a Personal Data Assistant (PDA), while the network 14 may comprise Virtual Private Network (VPN) accessed directly, or through an intermediate network (not shown). In the illustrated embodiment, the security mechanism 10 is a network card that provides VPN functionality and that has the configuration of a Personal Computer Memory Card International Association (PCMCIA) package for receipt in a PCMCIA slot 16 within the network peripheral device 12. Alternatively, the security mechanism 10 could take on other configurations and could offer different functionality without departing from the spirit and scope of the invention.

[0011] To facilitate the establishment of a secure session, the security mechanism 10 includes at least one immutable memory element 18 in the form of a Read Only Memory (ROM) element that stores information (software and support files) that remains fixed for all times (i.e., for each and every communications session). In the illustrated embodiment, the memory element 18 bears the designation "Security ROM" because it stores security application software (including bootstrap) code that initiates the operation of the security mechanism 10. Further, the security ROM 18 also generates at random the private key(s) required by the security mechanism 10 to perform its security functions including user and device authentication. (This ensures that the private keys used by different network peripheral devices remain independent from each other and that the security device cannot be forced to use keys known to an attacker.)

[0012] As an adjunct to Security ROM 18, the security mechanism 10 may also include a write-once ROM 20 for storing information written into the ROM during manufacture of the security mechanism. Such information may include additional bootstrap code as well as any upgrade that occurred subsequent to the manufacture of the Security ROM 18. (To the extent that either of the ROMs 18 and 20 have an upgrade capability, only the upgrade management software should have the capability of modifying the software in each ROM. Further, any application that loads either ROM should be signed by the manufacturer and the signature verified prior to writing any data.)

[0013] In addition to the security ROM 18 (and the write-once ROM 20 if present), the security mechanism 10 of FIG. 1 also includes at least one "persistent" memory element 24, in the form of a CMOS Random Access Memory (RAM) or a Programmable Read Only Memory (PROM) for receiving data prior to or during a communications session and for retaining such data for use during a subsequent session. In FIG. 1, the memory element 24 bears the designation "Configuration Memory" because this memory element stores configuration data that enables the security mechanism 10 to facilitate a connection with different networks. Thus, the contents of the Configuration memory element 24 can change upon an application executed by the network peripheral 12 that requires new or updated configuration information. To maintain security, only the application requiring new or updated configuration information should have the ability to write data to the configuration memory element 24 and the data written to this area must not be of a nature that could compromise the security afforded to the user. In other words, security critical data (i.e., data identifying the user and the device) must be excluded. The application executed by the network peripheral device 12 that seeks to write data to the Configuration Memory Element 24 should require signing and that such signing should be verified by the information in the Security ROM 18.

[0014] In addition to the previously described memory elements, the security mechanism 10 also includes at least one volatile memory element 26 in the form of a Random Access (RAM) memory or the like. The RAM 26 holds session-specific data, including user-entered verification data, such as a password or PIN, as well as authentication data generated by the security mechanism 10 itself. The data held within the RAM 26 remains only for the duration of a session. At the end of each session, as

well as upon a power-down condition, the bootstrap code within the Security ROM 18 (or the bootstrap code in the Write-Once ROM 20) causes the RAM 26 to erase all of its data (or at least its sensitive security data) if such data has not already been erased. In this way, the memory element 26 loses all user-entered verification data, as well as all security mechanism-generated authentication data associated with a given session upon its completion, or upon a power-down condition.

[0015] An interconnection medium 28 in the form of a circuit board or the like, supports and interconnects the Security ROM 18, the Write-once ROM 20, the Configuration 24 memory and the volatile memory 26, as well as other chips (not shown) such as a central processing unit. The circuit board 28 couples the memory elements and other components mounted thereon to a connector 30, which mates with a complementary connector (not shown) in the PCMCIA slot 16 of the network peripheral device 12. A tamper-evident enclosure 32 surrounds the circuit board 28 and the components mounted thereon to prevent access to such components, thus preventing tampering therewith. The effective level of the physical security of the security mechanism 10 depends the selection of the materials and fabrication technology employed. In addition to preventing access to the components on the circuit board 28, the tamper-evident enclosure 32 has the property that it readily exhibits any attempt to gain access there through to the circuit board and the components mounted thereon. In this way, a user who inspects the tamper-evident enclosure 32 can easily observe whether anyone has attempted to gain access to any of the Security ROM 18, the Write-once ROM 20, the Configuration 24 memory and the volatile memory 26, thereby compromising the integrity of the security mechanism 10. In addition to employing the tamper-evident enclosure 32, the components of the security mechanism 10 are strengthened against extreme environmental conditions, including, but not limited to under/over voltage conditions, fast/slow clock speeds, temperature variations and electromagnetic radiation.

[0016] In the illustrated preferred embodiment, the network security mechanism 10 implements VPN functionality using IPsec. At the beginning of each network connection session, the security mechanism 10 will obtain from the user the security critical data that describes the security services to be provided. This data should specify which security

gateway to connect to, which cryptographic algorithms and which key sizes are acceptable, the username by which the user is known to the security gateway and the password that the security gateway will use to authenticate the user. Using this security critical data, the security mechanism 10 establishes a secure connection to the indicated gateway, establishing encryption and authentication keys to be used for the remainder of the session as well as performing any authentication steps that are required by the security gateway to gain access to the resources it controls. The specifics of authentication and key establishment depend on the specific protocols in use. The details for the IPsec VPN protocol, for example, can be obtained from the definition of the protocol itself.

[0017] The above-described embodiments merely illustrate the principles of the invention. Those skilled in the art may make various modifications and changes that will embody the principles of the invention and fall within the spirit and scope thereof.